

# A UNIVERSAL BOUND ON THE VARIATIONS OF BOUNDED CONVEX FUNCTIONS

J. KWON

**ABSTRACT.** Given a convex set  $C$  in a real vector space  $E$  and two points  $x, y \in C$ , we investigate which are the possible values for the variation  $f(y) - f(x)$ , where  $f : C \rightarrow [m, M]$  is a bounded convex function. We then rewrite the bounds in terms of the Funk weak metric, which will imply that a bounded convex function is Lipschitz-continuous with respect to the Thompson and Hilbert metrics. The bounds are also proved to be optimal. We also exhibit the maximal subdifferential of a bounded convex function at a given point  $x \in C$ .

## 1. THE VARIATIONS OF BOUNDED CONVEX FUNCTIONS

Let  $C$  be a convex set of a real vector space  $E$ . Given two points  $x, y \in C$ , we define the following auxiliary quantity:

$$\tau_C(x, y) = \sup \{t \geq 1 \mid x + t(y - x) \in C\}.$$

Clearly,  $\tau_C$  takes values in  $[1, +\infty]$ . Intuitively, it measures how far away  $x$  is from the boundary in the direction of  $y$ , taking the “distance”  $xy$  as unit. Clearly,  $\tau_C(x, y) = +\infty$  if and only if  $x + \mathbb{R}_+(y - x) \subset C$ . Our first result is the following.

**Theorem 1.1.** *Let  $m \leq M$  be two real numbers. Let  $C$  be a convex set of a real vector space  $E$  and  $f : C \rightarrow [m, M]$  a convex function. For every couple of points  $(x, y) \in C^2$ ,  $f$  satisfies:*

$$-\frac{M - m}{\tau_C(y, x)} \leq f(y) - f(x) \leq \frac{M - m}{\tau_C(x, y)}.$$

*Proof.* It is enough to prove the result for functions with values in  $[0, 1]$ , since we can consider  $(M - m)^{-1}(f - m)$ . Let  $x, y$  be two points in  $C$ . Let  $t$  be such that  $1 \leq t < \tau_C(x, y)$ . By definition of  $\tau_C$ , and because  $C$  is convex, we have  $x + t(y - x) \in C$ . We can write  $y$  as a convex combination of  $x + t(y - x)$  and  $x$  with coefficients  $1/t$  and  $(t - 1)/t$  respectively:

$$y = \frac{x + t(y - x) + (t - 1)x}{t}.$$

---

*Date:* March 18, 2015.

*2010 Mathematics Subject Classification.* 26B25, 52A05.

*Key words and phrases.* Convex Functions, Variations, Funk Metric, Thompson Metric, Hilbert Metric.

The author is grateful to Pierre-Antoine Corre, Rida Laraki, Sylvain Sorin and Guillaume Vigeral for their very helpful comments.

By convexity of  $f$ , we get:

$$\begin{aligned} f(y) - f(x) &\leq \frac{f(x + t(y - x)) + (t - 1)f(x)}{t} - f(x) \\ &\leq \frac{f(x + t(y - x)) - f(x)}{t} \leq \frac{1}{t}, \end{aligned}$$

where the last inequality comes from the fact that  $f$  has values in  $[0, 1]$ . By taking the limit as  $t \rightarrow \tau_C(x, y)$ , we get:

$$f(y) - f(x) \leq \frac{1}{\tau_C(x, y)}.$$

The lower bound is obtained by exchanging the roles of  $x$  and  $y$ .  $\square$

## 2. THE FUNK, THOMPSON AND HILBERT METRICS

In this section, we rewrite the result from Theorem 1.1 as a Lipschitz-like property in the framework of convex sets in normed spaces. But  $1/\tau_C$  is far from being a distance. We thus consider the Funk, Thompson and Hilbert metrics (which were introduced in [1], [4] and [2] respectively) and establish the link with  $\tau_C$ .

We restrict our framework to the case where  $C$  is an open convex subset of a normed space  $(E, \|\cdot\|)$ . Let  $x, y \in C$ . If  $\tau_C(x, y) < +\infty$ , we can define  $b(x, y)$  to be the following point:

$$b(x, y) = x + \tau_C(x, y)(y - x).$$

Note that since  $C$  is open, when  $b(x, y)$  exists, it is necessarily different from  $y$ . This will be necessary to state the following definitions.

**Definition 2.1.** Let  $C$  be an open convex subset of a normed space  $(E, \|\cdot\|)$ . We define

(i) the Funk weak metric:

$$F_C(x, y) = \begin{cases} \log \frac{\|x - b(x, y)\|}{\|y - b(x, y)\|} & \text{if } \tau_C(x, y) < +\infty; \\ 0 & \text{otherwise} \end{cases};$$

(ii) the Thompson pseudometric:

$$T_C(x, y) = \max(F_C(x, y), F_C(y, x));$$

(iii) the Hilbert pseudometric:

$$H_C(x, y) = \frac{1}{2} (F_C(x, y) + F_C(y, x)).$$

**REMARK 2.2.** Even if we will abusively call them *metrics*, they fail to satisfy the separation axiom in general. The Thompson and the Hilbert metrics are thus *pseudometrics*. Moreover, the Funk metric not being symmetric, it actually is a *weak* metric. The Thompson and the Hilbert metrics are respectively the *max-symmetrization* and *meanvalue-symmetrisation* of the Funk metric. For a detailed presentation of these notions, see e.g. [3].

We now establish the link between  $\tau_C(x, y)$  and  $F_C(x, y)$ .

**Proposition 2.3.** *Let  $C$  be an open convex subset of a normed space  $(E, \|\cdot\|)$ . For every points  $x, y \in C$ , the following equality holds:*

$$F_C(x, y) = -\log \left( 1 - \frac{1}{\tau_C(x, y)} \right).$$

*Proof.* Let  $x, y \in C$ . If  $\tau_C(x, y) = +\infty$ , the right-hand side of the above equality is zero, as expected. If  $\tau_C(x, y) < +\infty$ ,  $\tau_C(x, y)$  can be expressed with the norm. Since by definition  $b(x, y) = x + \tau_C(x, y)(y - x)$ , we have

$$\tau_C(x, y) = \frac{\|x - b(x, y)\|}{\|x - y\|} \quad \text{and} \quad \tau_C(x, y) - 1 = \frac{\|y - b(x, y)\|}{\|x - y\|}.$$

And thus:

$$\frac{\|x - b(x, y)\|}{\|y - b(x, y)\|} = \left( 1 - \frac{1}{\tau_C(x, y)} \right)^{-1}.$$

Therefore,

$$F_C(x, y) = -\log \left( 1 - \frac{1}{\tau_C(x, y)} \right).$$

□

By combining Theorem 1.1 and the above proposition, we get the following corollary.

**Corollary 2.4.** *Let  $C$  an open convex subset of a normed space  $(E, \|\cdot\|)$  and  $f : C \rightarrow [m, M]$  be a convex function. Then, for all  $x, y \in C$ , the following bounds hold.*

- (i)  $-(M - m) \left( 1 - e^{-F_C(y, x)} \right) \leq f(y) - f(x) \leq (M - m) \left( 1 - e^{-F_C(x, y)} \right).$
- (ii)  $|f(y) - f(x)| \leq (M - m) \left( 1 - e^{-T_C(x, y)} \right).$
- (iii)  $|f(y) - f(x)| \leq (M - m) \left( 1 - e^{-2H_C(x, y)} \right).$

REMARK 2.5. From (ii), by using the inequality  $e^{-s} \geq 1 - s$ , we get:

$$\begin{aligned} |f(x) - f(y)| &\leq (M - m) \left( 1 - e^{-T_C(x, y)} \right) \\ &\leq (M - m) T_C(x, y), \end{aligned}$$

and similarly for (iii). Every convex function  $f : C \rightarrow [m, M]$  is thus  $(M - m)$ -Lipschitz (resp.  $2(M - m)$ -Lipschitz) with respect to the Thompson metric (resp. the Hilbert metric).

### 3. OPTIMALITY OF THE BOUNDS

We show in this section that the bounds obtained in Theorem 1.1 are optimal in the following sense. For a given convex set, and for a given couple a points, there is a function which attains the upper bound (resp. the lower bound). In other words, for  $x, y \in C$ :

$$\begin{cases} \max_{\substack{f: C \rightarrow [m, M] \\ f \text{ convex}}} (f(y) - f(x)) = \frac{M - m}{\tau_C(x, y)} \\ \min_{\substack{f: C \rightarrow [m, M] \\ f \text{ convex}}} (f(y) - f(x)) = -\frac{M - m}{\tau_C(y, x)}. \end{cases}$$

In the proof of the following theorem, it will be very convenient to extend the notion of convexity to functions defined on  $C$  and taking values in  $\mathbb{R} \cup \{-\infty\}$  (and not  $\mathbb{R} \cup \{+\infty\}$ ). Obviously, the result according to which the upper envelope of two convex functions is also a convex function remains true.

**Theorem 3.1.** *Let  $m \leq M$  be two real numbers. Let  $C$  be a convex set of a real vector space  $E$ . For every couple of points  $(x, y) \in C^2$ , there exists a convex function  $f : C \rightarrow [m, M]$  (resp.  $g : C \rightarrow [m, M]$ ) such that the upper bound (resp. lower bound) of Theorem 1.1 is attained; in other words:*

$$f(y) - f(x) = \frac{M - m}{\tau_C(x, y)} \quad \left( \text{resp. } g(y) - g(x) = -\frac{M - m}{\tau_C(y, x)} \right).$$

*Proof.* Let  $x$  and  $y$  be two points in  $C$ , and let us construct a convex function  $f : C \rightarrow [0, 1]$  satisfying the equality. If  $\tau_C(x, y) = +\infty$ , the bound is zero, and  $f = 0$  is adequate. From now on, we assume that  $\tau_C(x, y) < +\infty$ . The idea of the construction is the following. Let us first consider the line through  $x$  and  $y$ . We want  $f$  to increase from 0 at  $x$  to 1 at the boundary in the direction of  $y$ , in an affine way; and to be equal to zero in the other direction. Then, we will have to extend  $f$  to all  $C$  in a convex way. Let  $\vec{u} = \tau_C(x, y)(y - x)$ . For every  $z \in C$ , let us define  $\sigma(z) = \sup \{t \geq 0 \mid z + t\vec{u} \in C\}$ .  $\sigma$  clearly takes values in  $[0, +\infty]$ . Consider the following function.

$$\begin{aligned} \phi : C &\longrightarrow [-\infty, 1] \\ z &\longmapsto 1 - \sigma(z) \end{aligned}$$

Let us prove that  $\phi$  is convex. Let  $z_1$  and  $z_2$  be two points in  $C$  and  $z_3 = \lambda z_1 + (1 - \lambda)z_2$  (with  $\lambda \in (0, 1)$ ) a convex combination. By definition of  $\sigma$ , if we take two real numbers  $s_1$  and  $s_2$  such that  $0 \leq s_1 \leq \sigma(z_1)$  and  $0 \leq s_2 \leq \sigma(z_2)$ , we have:

$$\begin{cases} z_1 + s_1\vec{u} \in C \\ z_2 + s_2\vec{u} \in C. \end{cases}$$

And thus, the convex combination of these two points with coefficients  $\lambda$  and  $1 - \lambda$  also belongs to  $C$ :

$$\lambda(z_1 + s_1\vec{u}) + (1 - \lambda)(z_2 + s_2\vec{u}) \in C.$$

This point can be rewritten with  $z_3$ :

$$z_3 + (\lambda s_1 + (1 - \lambda)s_2)\vec{u} \in C.$$

By definition of  $\sigma(z_3)$ , we have  $\lambda s_1 + (1 - \lambda)s_2 \leq \sigma(z_3)$ . This inequality is true for every  $s_1 \leq \sigma(z_1)$  and  $s_2 \leq \sigma(z_2)$ . Consequently:

$$\lambda\sigma(z_1) + (1 - \lambda)\sigma(z_2) \leq \sigma(z_3).$$

We can now prove the convexity inequality.

$$\begin{aligned} \phi(z_3) = 1 - \sigma(z_3) &\leq 1 - (\lambda\sigma(z_1) + (1 - \lambda)\sigma(z_2)) \\ &= \lambda(1 - \sigma(z_1)) + (1 - \lambda)(1 - \sigma(z_2)) \\ &= \lambda\phi(z_1) + (1 - \lambda)\phi(z_2). \end{aligned}$$

We now choose  $f = \max(\phi, 0)$ . Since  $\phi \leq 1$ ,  $f$  takes values in  $[0, 1]$ . Let us prove that  $f$  satisfies the desired equality. Let us compute  $f(x)$  and  $f(y)$ .

$$\begin{aligned}\sigma(x) &= \sup \{t \geq 0 \mid x + t\vec{u} \in C\} \\ &= \sup \{t \geq 0 \mid x + t\tau_C(x, y)(y - x) \in C\} \\ &= \frac{1}{\tau_C(x, y)} \sup \{t' \geq 0 \mid x + t'(y - x) \in C\} \\ &= \frac{1}{\tau_C(x, y)} \tau_C(x, y) \\ &= 1.\end{aligned}$$

Thus  $\phi(x) = 1 - \sigma(x) = 0$  and  $f(x) = \max(0, 0) = 0$ . Similarly, we can prove:

$$\sigma(y) = \frac{\tau_C(x, y) - 1}{\tau_C(x, y)},$$

and thus,  $\phi(y) = 1 - \sigma(y) = \tau_C(x, y)^{-1}$  and  $f(y) = \max(\tau_C(x, y)^{-1}, 0) = \tau_C(x, y)^{-1}$ . We finally get:

$$f(y) - f(x) = \frac{1}{\tau_C(x, y)}.$$

The construction of  $g$  is analogous. □

#### 4. THE MAXIMAL SUBDIFFERENTIAL

In the case of a nonempty convex subset  $C \subset \mathbb{R}^n$ , and a given point  $x_0 \in C$ , we wonder what is the maximal subdifferential at  $x_0$  (in the sense of inclusion) for a function  $f : C \rightarrow [m, M]$ . We will prove that there *is* a maximal one, and will express it in terms of the subdifferential of a translation of the Minkowski gauge. For each  $x_0 \in C$ , we define  $g_{C, x_0} : C \rightarrow [0, 1]$  by

$$g_{C, x_0}(x) = \inf \{\lambda > 0 \mid x - x_0 \in \lambda(C - x_0)\}.$$

This function is obviously well-defined, and can be seen as a Minkowski gauge centered in  $x_0$  and restricted to  $C$ . It is well-known fact that the Minkowski gauge is a convex function. So is this one.

**Theorem 4.1.** *Let  $C$  be a nonempty convex subset of  $\mathbb{R}^n$  and  $x \in C$ . We have*

$$\max_{\substack{f: C \rightarrow [m, M] \\ f \text{ convex}}} \partial f(x) = (M - m) \partial g_{C, x}(x),$$

where the maximum is understood in the sense of inclusion.

*Proof.* Let us first relate  $g_{C, x_0}$  to  $\tau$ . Let  $x_0, x \in C$ . We have

$$\begin{aligned}g_{C, x_0}(x) &= \inf \{\lambda > 0 \mid x - x_0 \in \lambda(C - x_0)\} \\ &= \sup \left\{ t > 0 \mid x - x_0 \in \frac{1}{t}(C - x_0) \right\}^{-1} \\ &= \sup \{t > 0 \mid x_0 + t(x - x_0) \in C\}^{-1} \\ &= \frac{1}{\tau(x_0, x)}.\end{aligned}$$

Let us prove the result in the case  $m = 0$  and  $M = 1$ , from which the general case follows immediately. Let  $f : C \rightarrow [0, 1]$  be a convex function and  $x_0 \in C$ . Let us show that  $\partial f(x_0) \subset \partial g_{C, x_0}(x_0)$ . This is true if  $\partial f(x_0)$  is empty. Otherwise, let  $\zeta \in \partial f(x_0)$ . For every  $x \in C$ , we have

$$\begin{aligned} \langle \zeta | x - x_0 \rangle &\leq f(x) - f(x_0) \leq \frac{1}{\tau(x_0, x)} \\ &= g_{C, x_0}(x) = g_{C, x_0}(x) - g_{C, x_0}(x_0), \end{aligned}$$

where we used Theorem 1.1 for the second inequality. If  $x \notin C$ , the equality also holds, since  $g_{C, x_0}(x) = +\infty$ . We thus have  $\partial f(x_0) \subset \partial g_{C, x_0}(x_0)$ . We conclude by saying that  $g_{C, x_0}$  is a convex function on  $C$  with values in  $[0, 1]$ .  $\square$

#### REFERENCES

- [1] P. FUNK, *Über geometrien, bei denen die geraden die kürzesten sind*, Mathematische Annalen, 101 (1929), pp. 226–237.
- [2] D. HILBERT, *Über die gerade linie als kürzeste verbindung zweier punkte*, Mathematische Annalen, 46 (1895), pp. 91–96.
- [3] A. PAPADOPOULOS AND M. TROYANOV, *Weak metrics on Euclidean domains*, JP Journal of Geometry and Topology, 7 (2007), pp. 23–44.
- [4] A. C. THOMPSON, *On certain contraction mappings in a partially ordered vector space*, Proceedings of the American Mathematical Society, 14 (1963), pp. 438–443.

JOON KWON

INSTITUT DE MATHÉMATIQUES DE JUSSIEU  
ÉQUIPE COMBINATOIRE ET OPTIMISATION  
UNIVERSITÉ PIERRE-ET-MARIE-CURIE  
4 PLACE JUSSIEU  
75252 PARIS CEDEX 05 – FRANCE

E-MAIL: [joon.kwon@ens-lyon.org](mailto:joon.kwon@ens-lyon.org)